

УДК 81'33

А. А. Баркович, Н. С. Петрова

Минский государственный лингвистический университет

СЕНТИМЕНТ-АНАЛИЗ: ПРАГМАТИЧЕСКАЯ СПЕЦИФИКА

В статье рассмотрены прагматические аспекты практической реализации основных методик sentiment-анализа. Во введении очерчена специфика sentiment-анализа как актуального направления развития компьютерной лингвистики. В фокусе исследования – прагматика реализации методологического потенциала оценки тональности текста. В разделе, посвященном методам оценки тональности текста, дана характеристика приоритетов sentiment-анализа: целесообразности и результативности. Описаны методики sentiment-анализа и модели, лежащие в их основе. В разделе, посвященном результатам исследования и их оценке, отражена специфика использования методик тональных словарей, машинного обучения и машинного обучения с использованием тональных словарей. Иллюстративный материал представлен фрагментом англоязычного медиадискурса, репрезентативным для детализированного изучения уровневой специфики sentiment-анализа. Исследование базировалось на общепринятом категориальном фундаменте, в том числе охватывающем полярность, интенсивность и магнитуду тональности текста. Кроме традиционной лексико-грамматической парадигмы исследования текста, должное внимание было уделено таким инновационным лингвистическим категориям, как сущности и ключевые слова. Были идентифицированы такие категории метаданных, как бренды, геополитические реалии, национальности и ряд др.

В целом новизна проведенного исследования обеспечена лингвистическим анализом информационных процедур и методологической рефлексией интердисциплинарной проблематики. Не менее значимо рассмотрение прагматических аспектов деятельности по оценке тональности текста и актуализация в данной связи нового знания. Наиболее достоверные метаданные о тональности исследованного документа были получены с применением методики машинного обучения. Экспликация релевантных данных выполнялась программными средствами, специально предназначенными для проведения оценки тональности текста. Результаты данного исследования будут актуальны для оптимизации процедуры sentiment-анализа на всех ее этапах и имплементации соответствующих метаданных и знаний в лингвистическую и информационную практики.

Ключевые слова: sentiment-анализ, прагматика, методика, оценка тональности, тональный словарь, машинное обучение.

Для цитирования: Баркович А. А., Петрова Н. С. Sentiment-анализ: прагматическая специфика // Труды БГТУ. Сер. 4, Принт- и медиатехнологии. 2023. № 2 (273). С. 40–46. DOI: 10.52065/2520-6729-2023-273-2-6.

A. A. Barkovich, N. S. Petrova

Minsk State Linguistic University

SENTIMENT ANALYSIS: PRAGMATIC SPECIFICITY

The article considers the pragmatic aspects of the practical implementation of the main methods of sentiment analysis. The introduction outlines the specifics of sentiment analysis as an actual direction in the development of computational linguistics. The focus of the study is the pragmatics of implementing the methodological potential of the text tonality evaluation. In the section devoted to methods for the text tonality evaluation, a description is given of the priorities of sentiment analysis: expediency and effectiveness. The methods of sentiment analysis and the underlying models are described. The section devoted to the results of the study and their evaluation reflects the specifics of the use of methods of tone dictionaries, machine learning and machine learning using tone dictionaries. The illustrative material is presented by a fragment of the English-language media discourse, representative for a detailed study of the level specifics of sentiment analysis. The study was based on a generally accepted categorical foundation, including that covering the polarity, intensity and magnitude of the text tonality. In addition to the traditional lexical-grammatical paradigm of text research, due attention was paid to such innovative linguistic categories as entities and keywords. Such categories of metadata as brands, geopolitical realities, nationalities and a number of others were identified. The analysis is supported by detailed statistical data. In general, the novelty of this study is provided by the linguistic analysis of informational procedures and methodological reflection of interdisciplinary issues. Equally significant is the consideration of the

pragmalinguistic aspects of the text tonality evaluation and the actualization of new knowledge in this regard. In particular, a factually substantiated comparison of the main methods of sentiment analysis made it possible to draw meaningful conclusions about their advantages and drawbacks. The most reliable metadata about the tonality of the studied document was obtained using machine learning techniques. The explication of relevant data was carried out by software tools specially designed for evaluating the text tonality. The results of this study will be relevant for optimizing the sentiment analysis procedure at all its stages and implementing the relevant metadata and knowledge in linguistic and informational practices.

Keywords: sentiment analysis, pragmatics, methodology, text tonality, tone dictionary, machine learning.

For citation: Barkovich A. A., Petrova N. S. Sentiment analysis: pragmatic specificity. *Proceedings of BSTU, issue 4, Print- and Mediatechnologies*, 2023, no. 2 (273), pp. 40–46. DOI: 10.52065/2520-6729-2023-273-2-6 (In Russian).

Введение. Сентимент-анализ, также известный как оценка тональности текста, является быстро развивающейся технологией, которая использует достижения компьютерной лингвистики для выявления и извлечения субъективной информации из текстовых данных. История сентимент-анализа уходит своими корнями в 1950-е гг., когда соответствующий научный концепт был актуализирован исследователями в области психологии. Однако только в XXI в. термин *сентимент-анализ* стал широко известен и закрепился в научной парадигме: «Сентимент-анализ направлен на определение отношения говорящего или автора к некоторой теме или общей контекстуальной тональности документа» [1, с. 5]. Данное определение подчеркивает нацеленность сентимент-анализа на выявление в том или ином тексте эмоциональности [2]. Практическое применение сентимент-анализа широко и разнообразно, оно охватывает маркетинговые исследования и разработку продуктов, политический анализ и мониторинг социальных сетей. Одной из областей, в которой сентимент-анализ получил значительное распространение, является анализ новостных статей. Анализируя тональность, выраженную в новостных статьях, исследователи могут получить представление об общественном мнении и отношении читателей к тем или иным темам и событиям.

Основная часть. Методология исследования. Для достижения цели нашего исследования актуально рассмотрение релевантности не только логики программного инструментария, но и прагматически обоснованных методологических средств оценки тональности документа. Подобная универсальность подхода обеспечивает комплексность и объективность характеристики соответствующей деятельности. Для исследования объективной специфики сентимент-анализа безусловный интерес представляет прагматика его практической реализации. Именно те или иные прагматические установки предопределяют специфику осуществления любой деятельности. Интенциональные и оценочные аспекты

выполнения сентимент-анализа тесно коррелируют с его собственной прагматической обусловленностью как деятельности.

Естественным образом те или иные предпочтения полагающихся на сентимент-анализ пользователей коррелируют с категориями целесообразности и результативности. В данной связи они являются доминирующими. Приоритет *целесообразности* диктует ориентацию на использование релевантных средств оценки тональности текста. Приоритет *результативности* подразумевает достижимость поставленных перед сентимент-анализом задач. Данная прагматическая обусловленность предполагает уже рассмотрение актуального инструментария и критериев его выбора для решения той или иной задачи: «Выбор подходящего инструмента для сентимент-анализа играет критическую важную роль в получении точных результатов, и это зависит от специфики текста» [3, с. 2]. Для успешного проведения самого сентимент-анализа оптимальным является тот инструментарий, «... который наилучшим образом подходит для конкретных задач и исходных данных» [4, с. 639]. Вместе с тем задачи, решаемые в процессе сентимент-анализа, как правило, являются достаточно стереотипными. *Результатом* в такой деятельности считаются идентификация (стратегически) и квалификация (факультативно) эмоциональной семантики того либо иного документа. *Документом* в компьютерной лингвистике называется практически любой текст, даже если он состоит из одного предложения или синтаксически незавершенной речевой структуры – как это типичным образом происходит в современной интернет-коммуникации. Высокопрецизионные результаты сентимент-анализа – сложнодостижимая цель. Однако методология развивается и уже имеющиеся достижения позволяют говорить о сентимент-анализе как о продуктивном направлении компьютерной лингвистики. При этом рефлексия прагматических подходов к метаописанию практики анализа тональности текста остается дискуссионной и недостаточно проработанной.

Исследование сентимент-анализа как объекта будет репрезентативным и практически значимым при анализе референтного иллюстративного материала. В качестве такого материала может рассматриваться какой-либо текст, достаточно компактный и не слишком объемный, например статья новостного портала *The Washington Post* (от 20 октября 2022 г.), посвященная политике: *In the end, Liz Truss did not outlast a wilting lettuce* ('В конце концов, Лиз Трасс не смогла пережить увядающий салат-латук'). Данный текст находится в открытом доступе, написан литературным английским языком, состоит из 275 словоупотреблений и 1681 знака (с пробелами) (<https://www.washingtonpost.com>). Предполагается, что публицистический текст по умолчанию будет апеллировать к эмоциям своих потенциальных читателей. И, действительно, данный текст носит ироничный и критичный характер. При этом его ярко выраженная негативная тональность – вполне очевидный критерий верификации процедур его эмоциональной оценки посредством компьютерного инструментария.

В качестве инструментария сентимент-анализа могут быть задействованы находящиеся в открытом доступе программные ресурсы. Необходимо отметить, что к настоящему времени сформировалась соответствующая практика. Однако вся практика автоматизированной обработки естественного языка (NLP) остро нуждается в научно обоснованном осмыслении и регламентации, поскольку значительно опережает теоретическую рефлексию [5, с. 6]. Рациональные же принципы развития науки предполагают гармонизацию и синхронизацию теории и практики.

Практически на сегодняшний день реализуются три основные **методики** проведения сентимент-анализа:

- 1) методика тональных словарей;
- 2) методика машинного обучения;
- 3) методика машинного обучения с использованием тональных словарей.

Теоретически они в общих чертах соответствуют основным **моделям** автоматизированной обработки естественного языка:

- основанной на правилах модели;
- статистической модели;
- гибридной модели [6].

Существует множество классификаций соответствующих методик и моделей. Типична следующая точка зрения: «Мы идентифицируем методы обнаружения настроений как принадлежащие к одной из трех категорий, каждая из которых имеет свои достоинства и недостатки: словарные методы; методы машинного обучения с учителем и методы машинного обучения без учителя» [7, с. 2]. Достаточно популярна точка зрения, согласно которой выделяют мето-

дики машинного обучения, словарную и гибридную [8, 9]. Здесь, в частности, не упоминаются стоящие за ними модели, однако соответствующие им принципы так или иначе признаются актуальными в данной связи. Об этом в последнее время все чаще говорится прямо [10; 11]. И это целесообразно иметь в виду, поскольку, например, «Система анализа тональности ищет в рассматриваемом тексте слова, имеющие эмоционально-оценочный заряд, и, применяя заложенные в ней правила, учитывающие отрицание и слова-усилители, вычисляет тональность всего текста» [12, с. 1106]. Так или иначе, в словарях отражаются человеческие знания в виде правил, а машинное обучение, безусловно, статистически обеспечено. Совершенствование моделей и зависимых от них методик неизбежно приводит к появлению гибридных инструментов. Подобное развитие методологии присуще сфере компьютерной обработки естественного языка в целом; в том числе это заметно в области автоматизированного перевода [13].

Результаты исследования и их оценка. Для определения наиболее целесообразной и результативной методики необходимо охарактеризовать практику их задействования с учетом соответствующей прагматической интенциональности.

1. Методика тональных словарей может быть рассмотрена в контексте использования, например, инструментария *SenticNet* (<https://sentic.net>). Характерно, что данное программное обеспечение было создано не путем ручной обработки данных таких глобальных ресурсов, как *WordNet* или *DBPedia*. Вместо этого здесь используются уже структурированные метаданные, анализируются графы и масштабируются знания из таких источников, как *WordNet-Affect*, *Open Mind Common Sense* и *GECKA*. Соответствующие данные представлены на трех уровнях (в соответствии с концепцией паналогии М. Минского): семантическая сеть, матрица и векторное пространство [14, с. 303]. Формализация референтных семантических связей строится посредством «распространяющейся активации», «семантических нейронов» и «песочных часов эмоций». Так или иначе, инструментарий максимально оснащенный и актуализированный.

Фактически словарь тональности *SenticNet* присваивает значения тональности словам, фразам и предложениям на основе категориального аппарата полярности и интенсивности. *Полярность* характеризует положительность или отрицательность эмоциональной оценки. Численное выражение полярности колеблется от «-1» до «1». *Интенсивность* характеризует степень эмоциональной оценки. Численное выражение интенсивности колеблется от «0» до «1».

В результате проведения анализа в тональном словаре были обнаружены значения тональности для 81 словоупотребления из анализируемой статьи. Например, в используемом словаре нашли следующие прецеденты: *political* – «-0,837», *lettuce* – «0,237», *live stream* – «0,82», *wilting* – «-0,958» и т. д. Конечно, не все языковые единицы анализируемого документа оказались представлены в данном словаре. Тем не менее удалось сформировать его общую оценку тональности – «0,046».

Преимущества данной методики вполне очевидны:

а) простота применения: не требует компетентности на уровне эксперта (целесообразность);

б) быстрота использования: требуется лишь сопоставление словоупотреблений с заранее сформированной базой данных, тональным словарем (результативность).

Недостатки:

а) ограниченность словарного запаса: в частности, лишь 81 из 275 (29%) словоупотреблений статьи зафиксированы в тональном словаре;

б) игнорирование контекста: например, слово *lady* ('дама') считается позитивным, но словосочетание *The Iceberg Lady* ('дама-айсберг') в целом очевидно имеет негативную значимость;

в) игнорирование иронии и сарказма: например, значимость фразы *Lettuce Liz on Leaf Support* ('салат-латук Лиз с опорой для листьев') не будет оценена как отрицательная, хотя она явно содержит сарказм;

г) интерференция объектной области: например, в словосочетании *Lettuce Liz* лексема *lettuce*, по-видимому, означает глупого или некомпетентного человека (*When somebody is being extremely stupid or incompetent* – 'Когда кто-то ведет себя крайне глупо или некомпетентно') – как сленг имеет 440 вариантов значимости, очевидно не коррелируя с ботаникой [15];

д) субъективность словаря, зависящая от предпочтений его составителя;

е) возможная неактуальность данных и связанная с этим необходимость регулярной обновляемости словаря: он должен соответствовать хронологу речевой практики.

В целом релевантность методики тональных словарей неочевидна для решения сложных и специализированных задач.

2. Методика машинного обучения может реализовываться на базе программного обеспечения *Cloud Natural Language* на платформе *Google* (<https://cloud.google.com/natural-language>). Данный инструментарий основан на технологии облачных вычислений.

На основе доступной сервису статистики была сделана оценка тональности каждого из 14 предложений анализируемой статьи. Дополнительно

была оценена *магнитуда* – выраженность оценки вне зависимости от ее полярности. Анализ проводился с учетом дифференциации в тексте *сущностей*: в частности, были идентифицированы сущности "person" (например, *prime minister*), "organization" (например, *Daily Star*), "location" (например, *London*), "event" (например, *crisis*) и др. Соответственно, для предложения *Its market value soared amid a cost-of-living crisis and record inflation* ('Его рыночная стоимость взлетела на фоне кризиса стоимости жизни и рекордной инфляции') совпали численные значения интенсивности ("sentiment score") и магнитуды ("magnitude") – «0,2». Однако уже в предложении *It was the lettuce* ('Это был салат-латук') магнитуда оказалась «0,3», а интенсивность – «-0,3».

Таким образом, среднее значение магнитуды оценки тональности всей анализируемой статьи оказалось «5,7». При этом также были определены «категория» документа – *Новости* и его «тематика» – *Политика*. «Достоверность» текста получила оценку «0,88». Общая оценка тональности документа составила «-0,2» (отрицательная).

Преимущества данной методики следующие:

а) экономичность: методика машинного обучения позволяет автоматизировать процесс и использовать сетевые ресурсы (целесообразность);

б) разноаспектность: при наличии массива уже верифицированных данных программа может учитывать нюансы значимости, актуализируя предыдущий опыт оценки тональности тех или иных стереотипных фрагментов речевой продукции (результативность).

Недостатки:

а) игнорирование культурной специфики: статистический материал, как правило, не учитывает специфику определенной культуры или демографической группы, что может привести к ошибочным выводам при анализе культурно-специфичного материала;

б) возможная разнонаправленность, связанная с неоднородностью материала и необходимостью учитывать противоречивость частей документа – как в анализируемом документе.

В целом при определенной технологической усложненности процедуры машинного обучения ее релевантность обеспечена эффективностью.

3. Методика машинного обучения с использованием тональных словарей может базироваться, в частности, на программном обеспечении *Speak* (<https://speakai.co>). Данный сервис выполняет анализ речевых паттернов и оценивает тональность документа с учетом положительной, отрицательной или нейтральной полярности предварительно оцененных языковых единиц. При этом алгоритм использует машинное обучение для постоянного совершенствования процедуры анализа тональности текста.

Speak автоматически идентифицирует и анализирует аудио-, видео- и текстовые данные, выявляя ключевые слова, темы и шаблоны. Параметрами анализа являются полярность, ее интенсивность и эмоциональность. Эмоциональность подразумевает категории, в том числе *радости, печали, гнева, страха*. Алгоритмы сервиса настроены на учет юмора и иных контекстных особенностей документа. Функциональность *Speak* обеспечена совмещением методов автоматизированной обработки естественного языка и машинного обучения.

В результате задействования сервиса были выявлены такие категории метаданных, как бренды (“Brands”) – *YouTube, Daily Star*, актуальные для 15,2% тональности документа; геополитические реалии (“Geopolitical”) – *London, Britain, the United Kingdom*, актуальные для 9,1% тональности документа; национальности (“Nationalities”) – *British*, актуальные для 6,1% тональности документа; люди (“People”) – *Liz Truss, Vladimir Putin*, актуальные для 6,1% тональности документа. Также были идентифицированы ключевые слова, актуальные для 60,6% тональности документа, и установлена их частотность: *Truss* (4), *prime minister* (2), *Lettuce Liz* (1), *Iceberg Lady* (1), *leaf support* (1), *record inflation* (1).

Соответственно, для предложения *This year has been a great one for lettuce, at least in Britain* (‘Этот год был отличным для салата, по крайней мере, в Британии’) описание интенсивности оценки было квалифицировано как *очень удовлетворительное* (“very satisfied”) с числовым значением «0,6249». Для предложения же *Its market value soared amid a cost-of-living crisis and record inflation* (‘Его рыночная стоимость взлетела на фоне кризиса стоимости жизни и рекордной инфляции’) описание интенсивности оценки было квалифицировано как *неудовлетворительное* (“dissatisfied”) с числовым значением «-0,4019». В целом *Speak* квалифицировал оценку тональности 42,86% предложений как *положительную*, 21,43% – как *нейтральную* и 35,71% – как *негативную*. Также была автоматически сгенерирована диаграмма, отражающая данную статистику графически (рисунок).



Графическая интерпретация тональности анализируемого документа посредством *Speak*

Общая оценка тональности документа составила «0,0178» (положительная).

К *преимуществам* данной методики следует отнести:

а) надежность: задействование системы более обосновано, поскольку при этом используются разные инструментарию обработки данных (целесообразность);

б) точность: задействование и машинного обучения, и словарей тональности повышает общую результативность, поскольку обе методики дополняют друг друга (результативность);

Недостатки:

а) разнонаправленность компонентов общей системы: машинное обучение зависит от качества и количества данных, используемых для «обучения», в то время как словари тональности не учитывают контекст;

б) сложность: интеграция двух методик требует значительных технических знаний и навыков разработки программного обеспечения;

в) противоречивость: интерпретация данных разного происхождения влияет на их общую объективность.

В целом релевантность машинного обучения с применением тональных словарей позволяет ориентироваться на достоинства обеих самостоятельных методик, ее составляющих.

Общая оценка тональности проанализированного документа при использовании различных методик оказалась достаточно противоречивой. При этом выявлены как преимущества, так и недостатки использования разных моделей. Выполненная оценка тональности документа по методике тональных словарей составила «0,046». Аналогичный показатель при использовании методики машинного обучения составил «-0,2». И при задействовании методики машинного обучения с использованием тональных словарей результат – «0,0178». При общей очевидной негативной коннотации эмоциональности текста лишь методика машинного обучения позволила это подтвердить. Методики, связанные с использованием тональных словарей, не дали возможности зафиксировать изначально предсказуемый результат. Статистически опосредованный учет контекстуальной семантики, присущий методике машинного обучения, как видится, оказался весьма существенным фактором, повлиявшим на корректность сентимент-анализа.

Заключение. В статье дана характеристика прагматических приоритетов сентимент-анализа. Определяющей в данной связи является положительная семантика оценки, что позволяет ориентироваться на ее основные параметры-категории – целесообразность и результативность. Отрицательная семантика всегда вторична и, соответственно, более размыта.

Рамки метаописания «недостатков» как прагматической категории на практике оказались неопределенными и атрибутивно размытыми – ограниченность, игнорирование, интерференция, субъективность, неактуальность, сложность, разнонаправленность, противоречивость. Выявлена зависимость прагматических аспектов практической реализации основных методик сентимент-анализа от особенностей реализации методологического потенциала. Описаны методики сентимент-анализа и модели, лежащие в их основе. В разделе, посвященном результатам исследования и их оценке, отражена специфика применения методик тональных словарей, машинного обучения и машинного обучения с использованием тональных словарей. В качестве иллюстративного материала был использован фрагмент англоязычного медиадискурса, его репрезентативность оказалась достаточной для детализированного изучения уровневой специфики сентимент-анализа. Особенности оценки тональности текста рассматривались в контексте их полярности, интенсивности и магнитуды. Была идентифицирована практическая значи-

мость таких инновационных лингвистических категорий, как сущности и ключевые слова; проанализированы такие категории метаданных, как бренды, личности, геополитические реалии, события и ряд других. Анализ основывался на подробных статистических данных. В целом была реализована значимая новизна данного исследования, связанная с лингвистическим анализом информационных процедур и методологической рефлексией интердисциплинарной проблематики. Не менее значимы результаты рассмотрения прагмалингвистических аспектов деятельности по оценке тональности текста. Фактологически обоснованное сопоставление основных методик сентимент-анализа позволило сделать содержательные выводы об их достоинствах и недостатках. Наиболее достоверные метаданные о тональности исследованного документа были получены с использованием методики машинного обучения. Для проведения анализа были задействованы актуальные программные средства. Полученные результаты могут быть учтены в практике сентимент-анализа и использованы в описании его лингвоинформационной специфики.

Список литературы

1. Liu B., Hu M., Cheng J. Opinion Observer: Analyzing and Comparing Opinions on the Web // WWW–2005: Proceedings of the 14th International World Wide Web conference, May 10–14, 2005, Chiba, Japan. New York, 2005. P. 342–351.
2. Pang B., Lee L. Opinion Mining and Sentiment Analysis // Foundations and Trends in Information Retrieval. 2008. Vol. 2, no. 1–2. P. 1–135.
3. Feldman R. Techniques and applications for sentiment analysis // Communications of the ACM. 2013. Vol. 56, no. 4. P. 82–89.
4. Handbook of Natural Language Processing / eds.: N. Indurkha, F. J. Damerau. Boca Raton: Chapman & Hall / CRC, 2010. 702 p.
5. Barkovich A. Informational Linguistics: The New Communicational Reality Newcastle upon Tyne: Cambridge Scholars Publ., 2020. 271 p.
6. Баркович А. А. Интернет-дискурс: компьютерно-опосредованная коммуникация. М.: Флинта: Наука, 2015. 288 с.
7. Reagan A. J., Danforth C. M., Tivnan B., Williams J. R., Dodds P. S. Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs // EPJ Data Science. 2017, 6.
8. Poria S., Hazarika D., Majumder N., Mihalcea R. Beneath the tip of the iceberg: current challenges and new directions in sentiment analysis research // IEEE Transactions on Affective Computing. 2020. No. 14(1). P. 108–132.
9. Rana M. R. R., Nawaz A., Iqbal J. A survey on sentiment classification algorithms, challenges and applications // Acta Universitatis Sapientiae, Informatica. 2018. No. 10(1). P. 58–72.
10. Пазельская А. Г., Соловьев А. Н. Метод определения эмоций в текстах на русском языке // Компьютерная лингвистика и интеллектуальные технологии: ежегодная Междунар. конф. «Диалог», Бекасово, 25–29 мая 2011 г., М., 2011. Вып. 10. С. 510–522.
11. Taboada M., Brooke J., Tofiloski M., Voll K., Stede M. Lexicon-based methods for sentiment analysis // Computational linguistics. 2011. No. 37(2). P. 267–307.
12. Кулагин Д. И. Открытый тональный словарь русского языка КартаСловСент // Компьютерная лингвистика и интеллектуальные технологии: материалы ежегодной Междунар. конф. «Диалог», Москва, 16–19 июня 2021 г., М., 2021. Вып. 20. С. 1106–1119.
13. Баркович А. А. Тенденции развития компьютерно-опосредованного дискурса: металингвистический аспект // Вестник Тверского государственного университета. Филология. 2015, № 4. С. 108–114.

14. Minsky M. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York: Simon & Schuster, 2007. 400 p.
15. Urban Dictionary. URL: <https://www.urbandictionary.com/define.php?term=Lettuce> (accessed: 13.05.2023).

References

1. Liu B., Hu M., Cheng J. Opinion Observer: Analyzing and Comparing Opinions on the Web. *WWW-2005*. Proceedings of the 14th International World Wide Web conference, May 10–14, 2005, Chiba, Japan. New York. 2005, pp. 342–351.
2. Pang B., Lee L. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2008, vol. 2, no. 1–2, pp. 1–135.
3. Feldman R. Techniques and applications for sentiment analysis. *Communications of the ACM*. 2013, vol. 56, no. 4, pp. 82–89.
4. *Handbook of Natural Language Processing* / eds. N. Indurkha, F. J. Damerau. Boca Raton: Chapman & Hall / CRC, 2010. 702 p.
5. Barkovich A. *Informational Linguistics: The New Communicational Reality* Newcastle upon Tyne: Cambridge Scholars Publ., 2020. 271 p.
6. Barkovich A. A. *Internet-diskurs: komp'yuterno-oposredovannaya kommunikatsiya* [Internet discourse: computer-mediated communication]. Moscow, FLINTA Publ.; Nauka Publ., 2015. 288 p. (In Russian).
7. Reagan A. J., Danforth C. M., Tivnan B., Williams J. R., Dodds P. S. Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs. *EPJ Data Science*, 2017, 6.
8. Poria S., Hazarika D., Majumder N., Mihalcea R. Beneath the tip of the iceberg: current challenges and new directions in sentiment analysis research. *IEEE Transactions on Affective Computing*, 2020, no. 14(1), pp. 108–132.
9. Rana M. R. R., Nawaz A., Iqbal J. A survey on sentiment classification algorithms, challenges and applications. *Acta Universitatis Sapientiae, Informatica*, 2018, no. 10(1), pp. 58–72.
10. Pazel'skaya A. G., Solov'ev A. N. A method for determining emotions in texts in Russian. *Komp'yuternaya lingvistika i intellektual'nyye tekhnologii: ezhegodnaya Mezhdunarodnaya konferentsiya "Dialog"* [Computational Linguistics and Intelligent Technologies: annual International Conference "Dialogue"]. Moscow, 2011, issue 10, pp. 510–522 (In Russian).
11. Taboada M., Brooke J., Tofiloski M., Voll K., Stede M. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 2011, no. 37(2), pp. 267–307.
12. Kulagin D. I. Open Tonal Dictionary of the Russian Language KartaSlovSent. *Komp'yuternaya lingvistika i intellektual'nyye tekhnologii: materialy ezhegodnoy Mezhdunarodnoy konferentsii "Dialog"* [Computational Linguistics and Intelligent Technologies: materials of the annual International Conference "Dialogue"]. Moscow, 2021, issue 20, pp. 1106–1119 (In Russian).
13. Barkovich A. A. Trends in the development of computer-mediated discourse: metalinguistic aspect *Vestnik Tverskogo gosudarstvennogo universiteta* [Bulletin of the Tver State University], series Philology, 2015, no. 4, pp. 108–114 (In Russian).
14. Minsky M. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York: Simon & Schuster, 2007. 400 p.
15. Urban Dictionary. Available at: <https://www.urbandictionary.com/define.php?term=Lettuce> (accessed 13.05.2023).

Информация об авторах

Баркович Александр Аркадьевич – доктор филологических наук, доцент, заведующий кафедрой информатики и прикладной лингвистики. Минский государственный лингвистический университет (220034, г. Минск, ул. Захарова, 21, Республика Беларусь). E-mail: barkovichaa@gmail.com

Петрова Наталья Сергеевна – преподаватель. Минский государственный лингвистический университет (220034, г. Минск, ул. Захарова, 21, Республика Беларусь). E-mail: natali040100@gmail.com

Information about the authors

Barkovich Aleksandr Arkad'yevich – DSc (Philology), Associate Professor, Head of the Department of Informatics and Applied Linguistics. Minsk State Linguistic University (21, Zakharova str., 220034, Minsk, Republic of Belarus). E-mail: barkovichaa@gmail.com

Petrova Natal'ya Sergeevna – Lecturer. Minsk State Linguistic University (21, Zakharova str., 220034, Minsk, Republic of Belarus). E-mail: natali040100@gmail.com

Поступила 26.06.2023